



# Full wwPDB X-ray Structure Validation Report i

Oct 10, 2014 – 04:48 AM EDT

PDB ID : 4D2T  
Title : Structure of MELK in complex with inhibitors  
Authors : Johnson, C.N.; Berdini, V.; Beke, L.; Bonnet, P.; Brehmer, D.; Coyle, J.E.;  
Day, P.J.; Frederickson, M.; Freyne, E.J.E.; Gilissen, R.A.H.J.; Hamlett,  
C.C.F.; Howard, S.; Meerpoel, L.; McMenamin, R.; Patel, S.; Rees, D.C.;  
Sharff, A.; Sommen, F.; Wu, T.; Linders, J.T.M.; ,  
Deposited on : 2014-05-13  
Resolution : 2.70 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

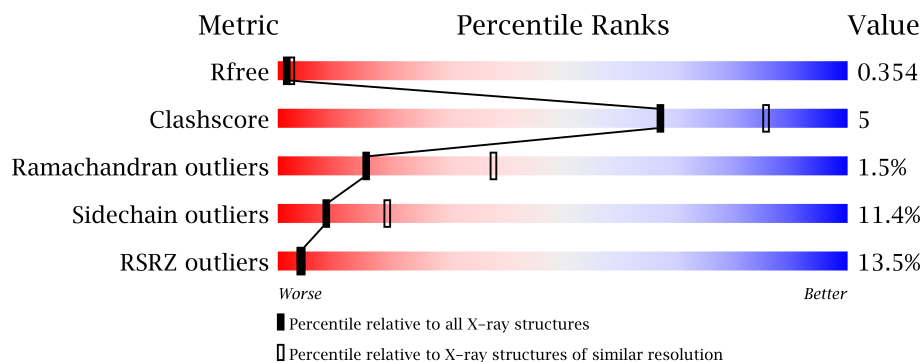
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.16 November 2013  
Xtriage (Phenix) : dev-1439  
EDS : stable23828  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable23828

# 1 Overall quality at a glance

The reported resolution of this entry is 2.70 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	66092	1557 (2.70-2.70)
Clashscore	79885	1939 (2.70-2.70)
Ramachandran outliers	78287	1905 (2.70-2.70)
Sidechain outliers	78261	1905 (2.70-2.70)
RSRZ outliers	66119	1559 (2.70-2.70)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	356	
1	B	356	
1	C	356	
1	D	356	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 10865 atoms, of which 84 are hydrogens and 0 are deuterium.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called MATERNAL EMBRYONIC LEUCINE ZIPPER KINASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	316	Total	C	N	O	S	0	0	0
			2554	1645	431	460	18			
1	B	323	Total	C	N	O	S	0	0	0
			2610	1677	444	471	18			
1	C	313	Total	C	N	O	S	0	1	0
			2544	1641	429	457	17			
1	D	319	Total	C	N	O	S	0	0	0
			2577	1666	430	465	16			

There are 108 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	EXPRESSION TAG	UNP Q14680
A	-18	GLY	-	EXPRESSION TAG	UNP Q14680
A	-17	SER	-	EXPRESSION TAG	UNP Q14680
A	-16	SER	-	EXPRESSION TAG	UNP Q14680
A	-15	HIS	-	EXPRESSION TAG	UNP Q14680
A	-14	HIS	-	EXPRESSION TAG	UNP Q14680
A	-13	HIS	-	EXPRESSION TAG	UNP Q14680
A	-12	HIS	-	EXPRESSION TAG	UNP Q14680
A	-11	HIS	-	EXPRESSION TAG	UNP Q14680
A	-10	HIS	-	EXPRESSION TAG	UNP Q14680
A	-9	SER	-	EXPRESSION TAG	UNP Q14680
A	-8	SER	-	EXPRESSION TAG	UNP Q14680
A	-7	GLY	-	EXPRESSION TAG	UNP Q14680
A	-6	LEU	-	EXPRESSION TAG	UNP Q14680
A	-5	VAL	-	EXPRESSION TAG	UNP Q14680
A	-4	PRO	-	EXPRESSION TAG	UNP Q14680
A	-3	ARG	-	EXPRESSION TAG	UNP Q14680
A	-2	GLY	-	EXPRESSION TAG	UNP Q14680
A	-1	SER	-	EXPRESSION TAG	UNP Q14680
A	0	HIS	-	EXPRESSION TAG	UNP Q14680
A	167	ALA	THR	ENGINEERED MUTATION	UNP Q14680

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	171	ALA	SER	ENGINEERED MUTATION	UNP Q14680
A	213	THR	ASN	ENGINEERED MUTATION	UNP Q14680
A	214	ALA	VAL	ENGINEERED MUTATION	UNP Q14680
A	215	ALA	MET	ENGINEERED MUTATION	UNP Q14680
A	218	VAL	TYR	ENGINEERED MUTATION	UNP Q14680
A	219	ALA	LYS	ENGINEERED MUTATION	UNP Q14680
B	-19	MET	-	EXPRESSION TAG	UNP Q14680
B	-18	GLY	-	EXPRESSION TAG	UNP Q14680
B	-17	SER	-	EXPRESSION TAG	UNP Q14680
B	-16	SER	-	EXPRESSION TAG	UNP Q14680
B	-15	HIS	-	EXPRESSION TAG	UNP Q14680
B	-14	HIS	-	EXPRESSION TAG	UNP Q14680
B	-13	HIS	-	EXPRESSION TAG	UNP Q14680
B	-12	HIS	-	EXPRESSION TAG	UNP Q14680
B	-11	HIS	-	EXPRESSION TAG	UNP Q14680
B	-10	HIS	-	EXPRESSION TAG	UNP Q14680
B	-9	SER	-	EXPRESSION TAG	UNP Q14680
B	-8	SER	-	EXPRESSION TAG	UNP Q14680
B	-7	GLY	-	EXPRESSION TAG	UNP Q14680
B	-6	LEU	-	EXPRESSION TAG	UNP Q14680
B	-5	VAL	-	EXPRESSION TAG	UNP Q14680
B	-4	PRO	-	EXPRESSION TAG	UNP Q14680
B	-3	ARG	-	EXPRESSION TAG	UNP Q14680
B	-2	GLY	-	EXPRESSION TAG	UNP Q14680
B	-1	SER	-	EXPRESSION TAG	UNP Q14680
B	0	HIS	-	EXPRESSION TAG	UNP Q14680
B	167	ALA	THR	ENGINEERED MUTATION	UNP Q14680
B	171	ALA	SER	ENGINEERED MUTATION	UNP Q14680
B	213	THR	ASN	ENGINEERED MUTATION	UNP Q14680
B	214	ALA	VAL	ENGINEERED MUTATION	UNP Q14680
B	215	ALA	MET	ENGINEERED MUTATION	UNP Q14680
B	218	VAL	TYR	ENGINEERED MUTATION	UNP Q14680
B	219	ALA	LYS	ENGINEERED MUTATION	UNP Q14680
C	-19	MET	-	EXPRESSION TAG	UNP Q14680
C	-18	GLY	-	EXPRESSION TAG	UNP Q14680
C	-17	SER	-	EXPRESSION TAG	UNP Q14680
C	-16	SER	-	EXPRESSION TAG	UNP Q14680
C	-15	HIS	-	EXPRESSION TAG	UNP Q14680
C	-14	HIS	-	EXPRESSION TAG	UNP Q14680
C	-13	HIS	-	EXPRESSION TAG	UNP Q14680
C	-12	HIS	-	EXPRESSION TAG	UNP Q14680
C	-11	HIS	-	EXPRESSION TAG	UNP Q14680

*Continued on next page...*

*Continued from previous page...*

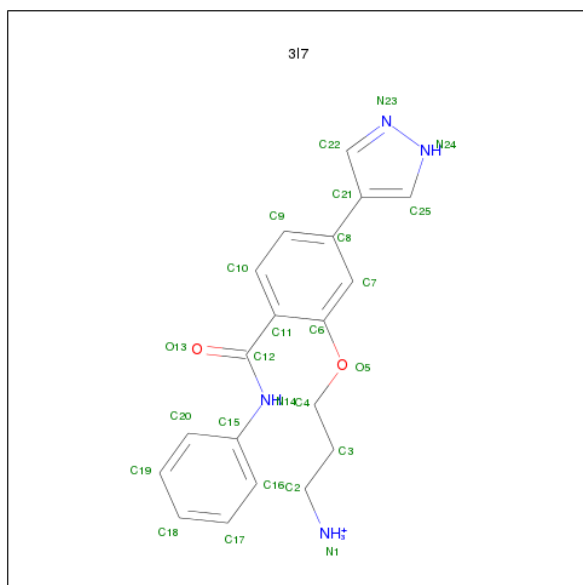
Chain	Residue	Modelled	Actual	Comment	Reference
C	-10	HIS	-	EXPRESSION TAG	UNP Q14680
C	-9	SER	-	EXPRESSION TAG	UNP Q14680
C	-8	SER	-	EXPRESSION TAG	UNP Q14680
C	-7	GLY	-	EXPRESSION TAG	UNP Q14680
C	-6	LEU	-	EXPRESSION TAG	UNP Q14680
C	-5	VAL	-	EXPRESSION TAG	UNP Q14680
C	-4	PRO	-	EXPRESSION TAG	UNP Q14680
C	-3	ARG	-	EXPRESSION TAG	UNP Q14680
C	-2	GLY	-	EXPRESSION TAG	UNP Q14680
C	-1	SER	-	EXPRESSION TAG	UNP Q14680
C	0	HIS	-	EXPRESSION TAG	UNP Q14680
C	167	ALA	THR	ENGINEERED MUTATION	UNP Q14680
C	171	ALA	SER	ENGINEERED MUTATION	UNP Q14680
C	213	THR	ASN	ENGINEERED MUTATION	UNP Q14680
C	214	ALA	VAL	ENGINEERED MUTATION	UNP Q14680
C	215	ALA	MET	ENGINEERED MUTATION	UNP Q14680
C	218	VAL	TYR	ENGINEERED MUTATION	UNP Q14680
C	219	ALA	LYS	ENGINEERED MUTATION	UNP Q14680
D	-19	MET	-	EXPRESSION TAG	UNP Q14680
D	-18	GLY	-	EXPRESSION TAG	UNP Q14680
D	-17	SER	-	EXPRESSION TAG	UNP Q14680
D	-16	SER	-	EXPRESSION TAG	UNP Q14680
D	-15	HIS	-	EXPRESSION TAG	UNP Q14680
D	-14	HIS	-	EXPRESSION TAG	UNP Q14680
D	-13	HIS	-	EXPRESSION TAG	UNP Q14680
D	-12	HIS	-	EXPRESSION TAG	UNP Q14680
D	-11	HIS	-	EXPRESSION TAG	UNP Q14680
D	-10	HIS	-	EXPRESSION TAG	UNP Q14680
D	-9	SER	-	EXPRESSION TAG	UNP Q14680
D	-8	SER	-	EXPRESSION TAG	UNP Q14680
D	-7	GLY	-	EXPRESSION TAG	UNP Q14680
D	-6	LEU	-	EXPRESSION TAG	UNP Q14680
D	-5	VAL	-	EXPRESSION TAG	UNP Q14680
D	-4	PRO	-	EXPRESSION TAG	UNP Q14680
D	-3	ARG	-	EXPRESSION TAG	UNP Q14680
D	-2	GLY	-	EXPRESSION TAG	UNP Q14680
D	-1	SER	-	EXPRESSION TAG	UNP Q14680
D	0	HIS	-	EXPRESSION TAG	UNP Q14680
D	167	ALA	THR	ENGINEERED MUTATION	UNP Q14680
D	171	ALA	SER	ENGINEERED MUTATION	UNP Q14680
D	213	THR	ASN	ENGINEERED MUTATION	UNP Q14680
D	214	ALA	VAL	ENGINEERED MUTATION	UNP Q14680

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	215	ALA	MET	ENGINEERED MUTATION	UNP Q14680
D	218	VAL	TYR	ENGINEERED MUTATION	UNP Q14680
D	219	ALA	LYS	ENGINEERED MUTATION	UNP Q14680

- Molecule 2 is 3-[2-(PHENYLCARBAMOYL)-5-(1H-PYRAZOL-4-YL)PHENOXY]PROPA N-1-AMINIUM (three-letter code: 3I7) (formula: C<sub>19</sub>H<sub>21</sub>N<sub>4</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	H	N	O	0	0
			46	19	21	4	2		
2	B	1	Total	C	H	N	O	0	0
			46	19	21	4	2		
2	C	1	Total	C	H	N	O	0	0
			46	19	21	4	2		
2	D	1	Total	C	H	N	O	0	0
			46	19	21	4	2		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	93	Total	O	0	0
			93	93		
3	B	126	Total	O	0	0
			126	126		
3	C	76	Total	O	0	0
			76	76		

Continued on next page...

*Continued from previous page...*

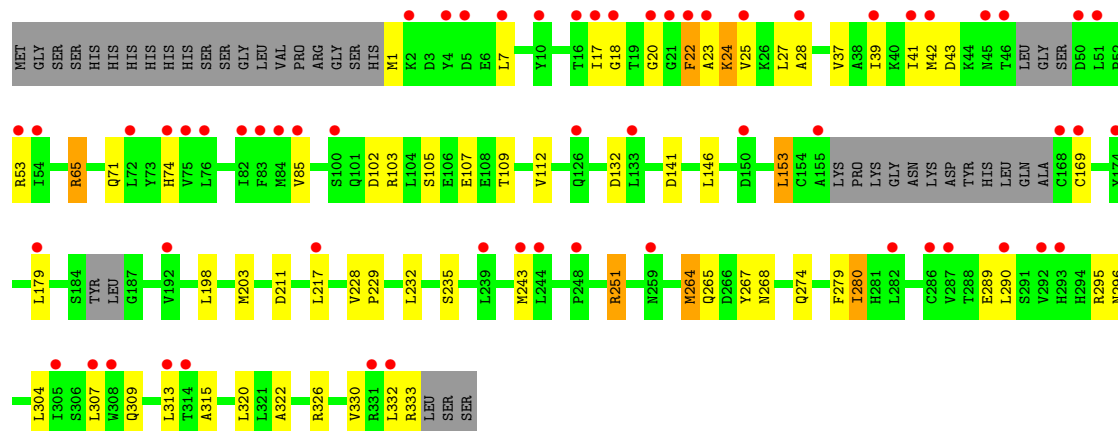
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	D	101	Total 101	O 101	0	0

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

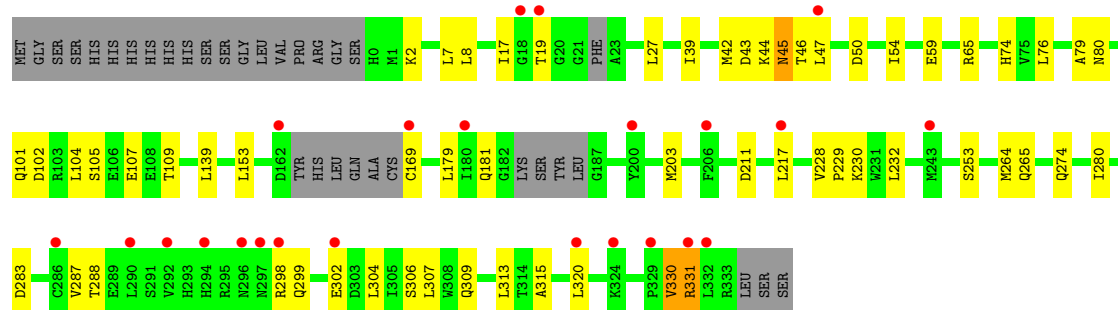
#### • Molecule 1: MATERNAL EMBRYONIC LEUCINE ZIPPER KINASE

Chain A: 



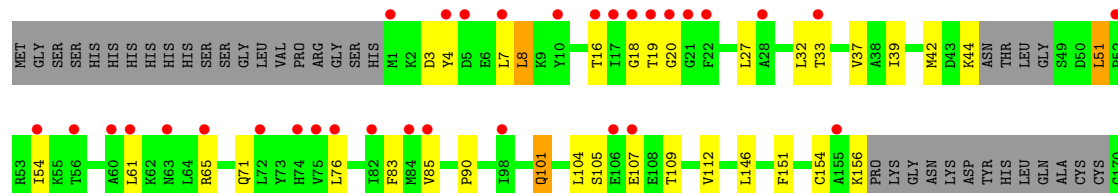
#### • Molecule 1: MATERNAL EMBRYONIC LEUCINE ZIPPER KINASE

Chain B: 

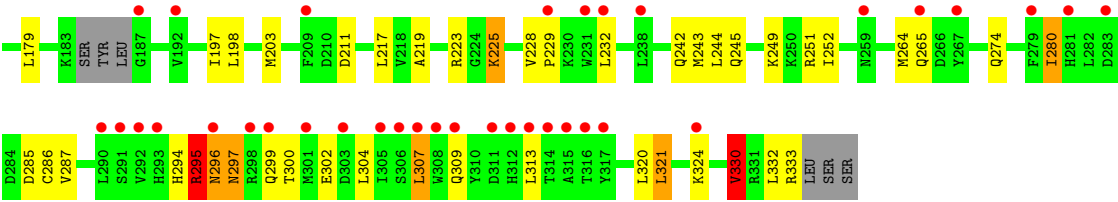


#### • Molecule 1: MATERNAL EMBRYONIC LEUCINE ZIPPER KINASE

Chain C: 

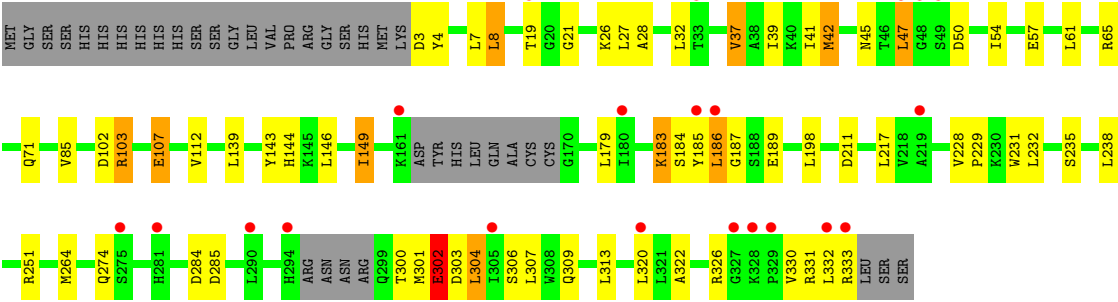






● Molecule 1: MATERNAL EMBRYONIC LEUCINE ZIPPER KINASE

Chain D:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.53Å 75.41Å 79.74Å 86.03° 69.05° 90.03°	Depositor
Resolution (Å)	48.54 – 2.70 48.54 – 2.70	Depositor EDS
% Data completeness (in resolution range)	93.1 (48.54-2.70) 93.1 (48.54-2.70)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.47 (at 2.69Å)	Xtriage
Refinement program	BUSTER 2.11.5	Depositor
R, $R_{free}$	0.270 , 0.340 0.274 , 0.354	Depositor DCC
$R_{free}$ test set	1870 reflections (5.34%)	DCC
Wilson B-factor (Å <sup>2</sup> )	71.0	Xtriage
Anisotropy	0.364	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 104.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.53$ , $\langle L^2 \rangle = 0.37$	Xtriage
Outliers	0 of 36893 reflections	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	10865	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	89.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.51% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 3I7

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# $ Z  > 5$	RMSZ	# $ Z  > 5$
1	A	0.51	0/2610	0.68	0/3529
1	B	0.52	0/2667	0.68	0/3603
1	C	0.49	0/2603	0.66	0/3517
1	D	0.52	0/2636	0.70	0/3565
All	All	0.51	0/10516	0.68	0/14214

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2554	0	2560	26	0
1	B	2610	0	2630	19	0
1	C	2544	0	2564	36	0
1	D	2577	0	2599	25	0
2	A	25	21	0	1	0
2	B	25	21	0	0	0
2	C	25	21	0	1	0
2	D	25	21	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	93	0	0	0	0
3	B	126	0	0	2	0
3	C	76	0	0	0	0
3	D	101	0	0	0	0
All	All	10781	84	10353	105	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 5.

All (105) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:302:GLU:HA	1:D:303:ASP:HB2	1.71	0.73
1:A:132:ASP:HB2	1:A:153:LEU:HD12	1.72	0.71
1:D:57:GLU:HG3	1:D:61:LEU:HD23	1.73	0.71
1:B:7:LEU:HD11	1:B:39:ILE:HD13	1.73	0.70
1:D:302:GLU:HB2	1:D:304:LEU:N	2.07	0.70
1:D:302:GLU:HB2	1:D:304:LEU:H	1.56	0.69
1:D:32:LEU:HD22	1:D:332:LEU:HG	1.74	0.69
1:A:25:VAL:HG21	2:A:1334:3I7:H3A	1.77	0.67
1:B:47:LEU:HB3	1:B:50:ASP:HB2	1.79	0.65
1:D:7:LEU:HD11	1:D:39:ILE:HD13	1.80	0.64
1:B:109:THR:HG21	1:B:203:MET:HG2	1.80	0.63
1:A:22:PHE:HB2	1:A:42:MET:HA	1.80	0.63
1:A:109:THR:HG21	1:A:203:MET:HG2	1.81	0.63
1:C:76:LEU:HB3	1:C:83:PHE:HB2	1.81	0.62
1:B:104:LEU:HB2	1:B:109:THR:HG22	1.79	0.62
1:A:264:MET:O	1:A:267:TYR:O	2.17	0.61
1:C:105:SER:O	1:C:109:THR:HG23	2.01	0.60
1:B:105:SER:O	1:B:109:THR:HG23	2.01	0.60
1:C:109:THR:HG21	1:C:203:MET:HG2	1.83	0.60
1:A:22:PHE:C	1:A:24:LYS:H	2.06	0.59
1:B:45:ASN:HD21	1:B:80:ASN:HB3	1.66	0.59
1:D:185:TYR:HD2	1:D:187:GLY:HA2	1.68	0.58
1:B:50:ASP:O	1:B:54:ILE:HG12	2.04	0.57
1:A:105:SER:O	1:A:109:THR:HG23	2.04	0.57
1:C:324:LYS:HD3	1:C:330:VAL:HG22	1.87	0.56
1:C:104:LEU:HB2	1:C:109:THR:HG22	1.87	0.56
1:C:243:MET:HG3	1:C:252:ILE:HD11	1.88	0.56
1:C:296:ASN:HD22	1:C:297:ASN:H	1.53	0.56
1:D:322:ALA:O	1:D:326:ARG:HG3	2.06	0.55
1:C:197:ILE:HD13	1:C:244:LEU:HD21	1.90	0.52
1:D:50:ASP:O	1:D:54:ILE:HG12	2.09	0.52

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:22:PHE:CD2	1:A:42:MET:HG3	2.44	0.52
1:C:39:ILE:HG12	1:C:85:VAL:HG22	1.93	0.51
1:D:28:ALA:HB3	1:D:37:VAL:HG12	1.92	0.51
1:D:320:LEU:HD22	1:D:330:VAL:HG13	1.93	0.51
1:C:219:ALA:HB1	1:D:107:GLU:HB2	1.93	0.51
1:A:18:GLY:HA3	1:A:25:VAL:H	1.77	0.50
1:C:242:GLN:HB3	1:C:252:ILE:HD13	1.93	0.50
1:C:242:GLN:HE21	1:C:252:ILE:HG21	1.75	0.50
1:B:288:THR:HB	1:B:298:ARG:HH12	1.76	0.50
1:C:61:LEU:HD11	1:C:151:PHE:CD1	2.46	0.50
1:C:229:PRO:HD2	1:C:232:LEU:HD22	1.94	0.50
1:B:283:ASP:O	1:B:287:VAL:HG23	2.12	0.49
1:A:65:ARG:HH21	1:A:279:PHE:HA	1.77	0.49
1:C:32:LEU:HB3	1:C:332:LEU:HD23	1.94	0.49
1:C:65:ARG:HG3	1:C:71:GLN:HE22	1.77	0.49
1:D:65:ARG:HG3	1:D:71:GLN:HE22	1.77	0.48
1:A:65:ARG:HG3	1:A:71:GLN:HE22	1.78	0.48
1:A:229:PRO:HD2	1:A:232:LEU:HD22	1.95	0.48
1:A:28:ALA:HB3	1:A:37:VAL:HG23	1.94	0.47
1:B:104:LEU:HB2	1:B:109:THR:CG2	2.45	0.47
1:D:229:PRO:HD2	1:D:232:LEU:HD22	1.96	0.47
1:A:320:LEU:HD22	1:A:330:VAL:HG13	1.96	0.47
1:D:143:TYR:OH	1:D:326:ARG:HG2	2.14	0.47
1:C:294:HIS:O	1:C:295:ARG:HB2	2.15	0.47
1:A:24:LYS:HD2	1:A:41:ILE:HB	1.97	0.47
1:C:280:ILE:H	1:C:280:ILE:HG13	1.59	0.47
1:A:289:GLU:HG3	1:A:332:LEU:HD22	1.97	0.46
1:C:243:MET:O	1:C:251:ARG:HD2	2.15	0.46
1:B:153:LEU:HD22	1:B:169:CYS:HB2	1.98	0.46
1:B:229:PRO:HD2	1:B:232:LEU:HD22	1.96	0.46
1:C:90:PRO:HA	2:C:1334:3I7:C17	2.47	0.45
1:C:104:LEU:HB2	1:C:109:THR:CG2	2.46	0.45
1:C:286:CYS:SG	1:C:321:LEU:HD12	2.57	0.45
1:C:324:LYS:HB2	1:C:330:VAL:HG13	2.00	0.44
1:B:307:LEU:HB3	1:B:309:GLN:HG3	1.99	0.44
1:C:42:MET:HG2	1:C:54:ILE:HG13	1.99	0.44
1:C:307:LEU:HB3	1:C:309:GLN:HG3	2.00	0.44
1:C:101:GLN:HG2	1:C:101:GLN:H	1.65	0.44
1:B:320:LEU:HD22	1:B:330:VAL:HG13	1.99	0.43
1:C:285:ASP:HB3	1:C:330:VAL:HG21	2.00	0.43
1:A:132:ASP:HB2	1:A:153:LEU:CD1	2.46	0.43
1:C:297:ASN:HD22	1:C:299:GLN:HB3	1.83	0.43

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:39:ILE:HG12	1:D:85:VAL:HG22	2.01	0.43
1:A:74:HIS:HB2	1:A:315:ALA:HB2	2.00	0.43
1:B:253:SER:HB2	3:B:2100:HOH:O	2.18	0.43
1:B:2:LYS:HG2	3:B:2003:HOH:O	2.18	0.43
1:C:33:THR:HG22	1:C:320:LEU:HG	2.01	0.43
1:D:183:LYS:O	1:D:185:TYR:HD1	2.01	0.43
1:A:22:PHE:CG	1:A:42:MET:HG3	2.54	0.42
1:C:7:LEU:HD11	1:C:39:ILE:HD13	2.01	0.42
1:A:280:ILE:HG13	1:A:280:ILE:H	1.62	0.42
1:B:331:ARG:HH11	1:B:331:ARG:HA	1.84	0.42
1:C:242:GLN:NE2	1:C:252:ILE:HG21	2.34	0.42
1:A:65:ARG:HB2	1:A:279:PHE:CE2	2.55	0.42
1:B:44:LYS:HE3	1:B:79:ALA:O	2.19	0.42
1:D:42:MET:HB3	1:D:47:LEU:HD13	2.02	0.41
1:A:112:VAL:HG13	1:A:146:LEU:HD11	2.03	0.41
1:A:322:ALA:O	1:A:326:ARG:HG3	2.20	0.41
1:D:307:LEU:HB3	1:D:309:GLN:HG3	2.02	0.41
1:C:4:TYR:HB3	1:C:8:LEU:HD12	2.03	0.41
1:D:139:LEU:HD12	1:D:149:ILE:CD1	2.51	0.41
1:A:39:ILE:HG12	1:A:85:VAL:HG22	2.02	0.41
1:D:4:TYR:HB3	1:D:8:LEU:HD12	2.03	0.41
1:C:243:MET:CG	1:C:252:ILE:HD11	2.51	0.41
1:A:243:MET:O	1:A:251:ARG:HD2	2.21	0.41
1:D:112:VAL:HG13	1:D:146:LEU:HD11	2.03	0.41
1:C:287:VAL:HG11	1:C:302:GLU:HG2	2.02	0.40
1:D:103:ARG:HD3	1:D:231:TRP:CE2	2.56	0.40
1:D:186:LEU:HB3	1:D:189:GLU:OE2	2.21	0.40
1:C:112:VAL:HG13	1:C:146:LEU:HD11	2.04	0.40
1:A:307:LEU:HB3	1:A:309:GLN:HG3	2.03	0.40
1:C:197:ILE:CD1	1:C:244:LEU:HD21	2.51	0.40
1:B:74:HIS:HB2	1:B:315:ALA:HB2	2.02	0.40
1:D:26:LYS:HE2	1:D:41:ILE:HD13	2.04	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	308/356 (86%)	285 (92%)	21 (7%)	2 (1%)	33	66
1	B	315/356 (88%)	296 (94%)	17 (5%)	2 (1%)	33	66
1	C	306/356 (86%)	280 (92%)	19 (6%)	7 (2%)	10	24
1	D	313/356 (88%)	285 (91%)	21 (7%)	7 (2%)	10	25
All	All	1242/1424 (87%)	1146 (92%)	78 (6%)	18 (1%)	15	41

All (18) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	19	THR
1	C	295	ARG
1	D	183	LYS
1	D	302	GLU
1	A	23	ALA
1	D	21	GLY
1	D	184	SER
1	C	265	GLN
1	C	330	VAL
1	D	304	LEU
1	C	225	LYS
1	D	306	SER
1	B	330	VAL
1	C	18	GLY
1	D	19	THR
1	C	51	LEU
1	A	20	GLY
1	C	20	GLY

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	277/314 (88%)	244 (88%)	33 (12%)	8	18
1	B	285/314 (91%)	255 (90%)	30 (10%)	10	22

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	277/314 (88%)	243 (88%)	34 (12%)	7	17
1	D	281/314 (90%)	250 (89%)	31 (11%)	9	21
All	All	1120/1256 (89%)	992 (89%)	128 (11%)	8	19

All (128) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	MET
1	A	7	LEU
1	A	17	ILE
1	A	22	PHE
1	A	24	LYS
1	A	27	LEU
1	A	43	ASP
1	A	53	ARG
1	A	65	ARG
1	A	102	ASP
1	A	103	ARG
1	A	107	GLU
1	A	141	ASP
1	A	153	LEU
1	A	169	CYS
1	A	179	LEU
1	A	198	LEU
1	A	211	ASP
1	A	217	LEU
1	A	228	VAL
1	A	235	SER
1	A	251	ARG
1	A	264	MET
1	A	265	GLN
1	A	268	ASN
1	A	274	GLN
1	A	280	ILE
1	A	290	LEU
1	A	295	ARG
1	A	296	ASN
1	A	304	LEU
1	A	313	LEU
1	A	333	ARG
1	B	8	LEU

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	B	17	ILE
1	B	27	LEU
1	B	42	MET
1	B	43	ASP
1	B	45	ASN
1	B	46	THR
1	B	59	GLU
1	B	65	ARG
1	B	76	LEU
1	B	101	GLN
1	B	102	ASP
1	B	107	GLU
1	B	139	LEU
1	B	179	LEU
1	B	181	GLN
1	B	211	ASP
1	B	217	LEU
1	B	228	VAL
1	B	230	LYS
1	B	264	MET
1	B	265	GLN
1	B	274	GLN
1	B	280	ILE
1	B	299	GLN
1	B	302	GLU
1	B	304	LEU
1	B	306	SER
1	B	313	LEU
1	B	331	ARG
1	C	3	ASP
1	C	8	LEU
1	C	16	THR
1	C	19	THR
1	C	27	LEU
1	C	37	VAL
1	C	44	LYS
1	C	51	LEU
1	C	101	GLN
1	C	107	GLU
1	C	154	CYS
1	C	156	LYS
1	C	179	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	198	LEU
1	C	211	ASP
1	C	217	LEU
1	C	223	ARG
1	C	225	LYS
1	C	228	VAL
1	C	245	GLN
1	C	249	LYS
1	C	264	MET
1	C	274	GLN
1	C	280	ILE
1	C	295	ARG
1	C	296	ASN
1	C	297	ASN
1	C	300	THR
1	C	304	LEU
1	C	307	LEU
1	C	313	LEU
1	C	321	LEU
1	C	330	VAL
1	C	333	ARG
1	D	3	ASP
1	D	8	LEU
1	D	27	LEU
1	D	37	VAL
1	D	42	MET
1	D	45	ASN
1	D	47	LEU
1	D	102	ASP
1	D	103	ARG
1	D	107	GLU
1	D	144	HIS
1	D	149	ILE
1	D	179	LEU
1	D	186	LEU
1	D	198	LEU
1	D	211	ASP
1	D	217	LEU
1	D	228	VAL
1	D	235	SER
1	D	238	LEU
1	D	251	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	264	MET
1	D	274	GLN
1	D	284	ASP
1	D	285	ASP
1	D	300	THR
1	D	301	MET
1	D	302	GLU
1	D	313	LEU
1	D	331	ARG
1	D	333	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (26) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	71	GLN
1	A	137	ASN
1	A	181	GLN
1	A	277	ASN
1	A	309	GLN
1	A	312	HIS
1	B	45	ASN
1	B	137	ASN
1	B	245	GLN
1	B	277	ASN
1	B	299	GLN
1	B	312	HIS
1	C	71	GLN
1	C	137	ASN
1	C	241	GLN
1	C	242	GLN
1	C	245	GLN
1	C	260	HIS
1	C	296	ASN
1	C	297	ASN
1	C	312	HIS
1	D	71	GLN
1	D	137	ASN
1	D	245	GLN
1	D	277	ASN
1	D	312	HIS

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

4 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	3I7	A	1334	-	27,27,27	0.49	0	35,35,35	0.98	2 (5%)
2	3I7	B	1334	-	27,27,27	0.65	1 (3%)	35,35,35	1.07	2 (5%)
2	3I7	C	1334	-	27,27,27	0.52	0	35,35,35	0.89	2 (5%)
2	3I7	D	1334	-	27,27,27	0.59	1 (3%)	35,35,35	1.05	2 (5%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	3I7	A	1334	-	-	0/17/17/17	0/3/3/3
2	3I7	B	1334	-	-	0/17/17/17	0/3/3/3
2	3I7	C	1334	-	-	0/17/17/17	0/3/3/3
2	3I7	D	1334	-	-	0/17/17/17	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1334	3I7	C25-C21	-2.23	1.38	1.39
2	D	1334	3I7	C25-C21	-2.11	1.38	1.39

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	1334	3I7	C21-C25-N24	-4.26	106.61	108.15
2	A	1334	3I7	C21-C25-N24	-4.06	106.68	108.15
2	B	1334	3I7	C21-C25-N24	-4.05	106.69	108.15
2	C	1334	3I7	C21-C25-N24	-3.14	107.01	108.15
2	B	1334	3I7	C21-C22-N23	-2.81	111.41	112.48
2	D	1334	3I7	C21-C22-N23	-2.51	111.53	112.48
2	C	1334	3I7	C21-C22-N23	-2.31	111.60	112.48
2	A	1334	3I7	C21-C22-N23	-2.16	111.66	112.48

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	316/356 (88%)	0.98	60 (18%) 2 2	47, 90, 135, 151	0
1	B	323/356 (90%)	0.56	23 (7%) 16 17	44, 80, 116, 139	0
1	C	313/356 (87%)	1.23	67 (21%) 1 2	53, 101, 142, 154	0
1	D	319/356 (89%)	0.51	21 (6%) 18 20	46, 82, 127, 164	0
All	All	1271/1424 (89%)	0.82	171 (13%) 4 4	44, 88, 132, 164	0

All (171) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	307	LEU	9.7
1	C	18	GLY	6.9
1	A	23	ALA	6.6
1	A	168	CYS	6.4
1	A	4	TYR	6.3
1	A	82	ILE	5.5
1	C	303	ASP	5.4
1	C	290	LEU	5.4
1	D	332	LEU	5.4
1	C	10	TYR	5.2
1	C	74	HIS	4.9
1	C	84	MET	4.9
1	C	308	TRP	4.7
1	A	292	VAL	4.6
1	A	21	GLY	4.5
1	C	5	ASP	4.4
1	C	305	ILE	4.4
1	A	53	ARG	4.3
1	A	75	VAL	4.3
1	C	313	LEU	4.3
1	A	5	ASP	4.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	267	TYR	4.1
1	A	18	GLY	4.0
1	B	18	GLY	4.0
1	B	162	ASP	4.0
1	C	292	VAL	4.0
1	C	309	GLN	3.9
1	C	293	HIS	3.8
1	C	298	ARG	3.8
1	A	282	LEU	3.8
1	A	84	MET	3.8
1	D	328	LYS	3.8
1	A	7	LEU	3.7
1	C	75	VAL	3.7
1	A	314	THR	3.6
1	C	155	ALA	3.6
1	C	21	GLY	3.6
1	C	60	ALA	3.5
1	C	229	PRO	3.5
1	D	281	HIS	3.5
1	C	301	MET	3.5
1	D	186	LEU	3.4
1	A	287	VAL	3.4
1	A	179	LEU	3.4
1	A	239	LEU	3.4
1	A	76	LEU	3.4
1	A	290	LEU	3.3
1	A	42	MET	3.3
1	C	259	ASN	3.3
1	B	180	ILE	3.3
1	C	82	ILE	3.3
1	C	315	ALA	3.2
1	B	296	ASN	3.2
1	A	192	VAL	3.2
1	C	72	LEU	3.2
1	C	324	LYS	3.2
1	C	107	GLU	3.2
1	A	22	PHE	3.1
1	A	169	CYS	3.1
1	A	83	PHE	3.1
1	A	313	LEU	3.1
1	D	19	THR	3.1
1	C	61	LEU	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	155	ALA	3.0
1	A	46	THR	3.0
1	A	2	LYS	3.0
1	A	286	CYS	3.0
1	C	85	VAL	2.9
1	D	305	ILE	2.9
1	D	290	LEU	2.9
1	A	305	ILE	2.9
1	D	320	LEU	2.9
1	C	1	MET	2.9
1	A	85	VAL	2.9
1	A	248	PRO	2.9
1	B	332	LEU	2.8
1	C	232	LEU	2.8
1	C	56	THR	2.8
1	B	324	LYS	2.8
1	D	180	ILE	2.7
1	A	10	TYR	2.7
1	A	293	HIS	2.7
1	C	306	SER	2.7
1	C	22	PHE	2.7
1	B	217	LEU	2.7
1	A	17	ILE	2.7
1	D	33	THR	2.7
1	A	50	ASP	2.6
1	A	244	LEU	2.6
1	C	52	PRO	2.6
1	C	283	ASP	2.6
1	B	169	CYS	2.6
1	C	299	GLN	2.6
1	B	294	HIS	2.6
1	B	206	PHE	2.6
1	A	150	ASP	2.6
1	A	20	GLY	2.5
1	B	331	ARG	2.5
1	C	187	GLY	2.5
1	A	45	ASN	2.5
1	A	72	LEU	2.5
1	A	100	SER	2.5
1	B	19	THR	2.5
1	C	28	ALA	2.5
1	C	238	LEU	2.5

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	311	ASP	2.5
1	B	297	ASN	2.4
1	A	307	LEU	2.4
1	A	259	ASN	2.4
1	C	19	THR	2.4
1	B	292	VAL	2.4
1	C	4	TYR	2.4
1	C	20	GLY	2.4
1	A	39	ILE	2.4
1	B	290	LEU	2.4
1	D	294	HIS	2.4
1	A	74	HIS	2.4
1	D	48	GLY	2.4
1	D	219	ALA	2.3
1	C	281	HIS	2.3
1	C	312	HIS	2.3
1	A	308	TRP	2.3
1	A	41	ILE	2.3
1	C	98	ILE	2.3
1	C	296	ASN	2.3
1	A	51	LEU	2.3
1	C	316	THR	2.3
1	A	54	ILE	2.3
1	C	17	ILE	2.3
1	B	243	MET	2.3
1	A	25	VAL	2.3
1	A	217	LEU	2.3
1	D	49	SER	2.3
1	B	200	TYR	2.2
1	A	133	LEU	2.2
1	C	314	THR	2.2
1	C	16	THR	2.2
1	C	7	LEU	2.2
1	C	54	ILE	2.2
1	A	174	TYR	2.2
1	D	333	ARG	2.2
1	D	161	LYS	2.2
1	D	327	GLY	2.2
1	A	16	THR	2.1
1	C	63	ASN	2.1
1	C	265	GLN	2.1
1	A	332	LEU	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	47	LEU	2.1
1	C	76	LEU	2.1
1	A	331	ARG	2.1
1	C	291	SER	2.1
1	C	106	GLU	2.1
1	C	192	VAL	2.1
1	C	317	TYR	2.1
1	D	329	PRO	2.1
1	C	65	ARG	2.1
1	C	231	TRP	2.1
1	C	279	PHE	2.1
1	D	275	SER	2.1
1	A	126	GLN	2.1
1	B	286	CYS	2.1
1	B	329	PRO	2.1
1	D	47	LEU	2.1
1	B	298	ARG	2.0
1	B	320	LEU	2.0
1	B	302	GLU	2.0
1	A	243	MET	2.0
1	D	185	TYR	2.0
1	C	33	THR	2.0
1	A	28	ALA	2.0
1	C	209	PHE	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q < 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	3I7	C	1334	25/25	0.38	1.00	64,76,93,93	46
2	3I7	A	1334	25/25	0.30	0.40	68,86,98,98	46
2	3I7	D	1334	25/25	0.20	-0.08	49,83,94,95	0
2	3I7	B	1334	25/25	0.21	-0.29	49,76,94,94	0

## 6.5 Other polymers

There are no such residues in this entry.